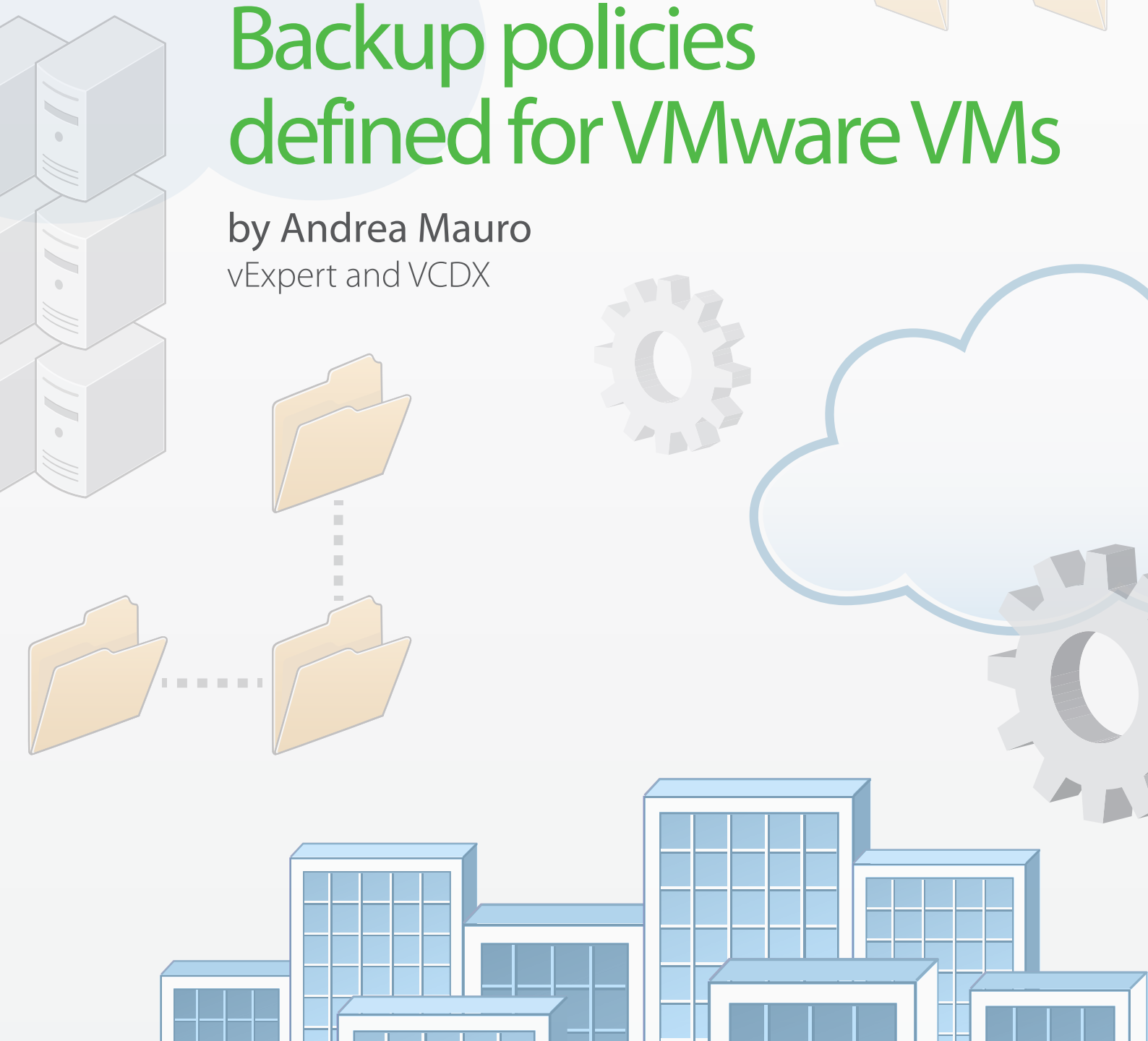
BACKUP Academy

Backup policies defined for VMware VMs

by Andrea Mauro
vExpert and VCDX



Author's bio

Andrea Mauro has worked in IT since 1995 and has several certifications (vExpert 2010/2011/2012, VCP, VCDX, MCITP, CCA and others). Andrea is a virtualization and storage architect, specializing in VMware (but also Microsoft, Citrix and Linux) solutions. Andrea is CIO of Assyrus Srl (http://www.assyrus.it), an Italian IT company founded in 2000.

Andrea is also a board member and founder of the VMware Italian User Group, VMTN Community Moderator, passionate blogger, runs his personal blog http://vinfrastructure.it (both in Italian and English).

Overview of backup policies

The challenge we have had over the years is to define what a backup policy is. Simple questions like what, where, how much, how often and others need to be reassessed for the virtualized era that we live in today. The properties of a backup policy are made up of critical details, including:

- Today, VMs include many different sources for a backup policy. These are files, application data, entire VM images and more.
- There also can be different types of strategies for the source of VMs. These are traditional full backups, incremental technologies and deduplication.
- From the management perspective, there are a lot of considerations around agent-based or agentless backups as well as around the way how the data is to be moved.
- In terms of a backup target, backup to disk, backup to tape, multi-level backup and more are among today's options for backup targets.

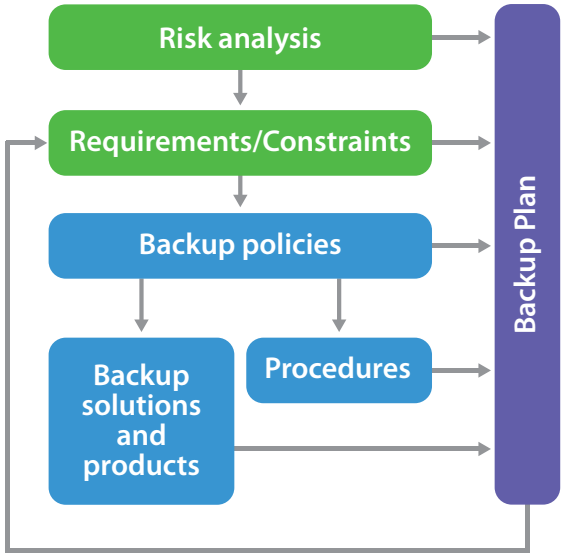
Finally, in this Backup Academy paper we will consider some aspects of backup retention, especially how virtualization has changed our data profile. We will discuss many considerations around retention vs. space and how different backup technologies impact the options associated with each strategy.

What is a backup policy?

A backup policy too often is based on technologies, practices and requirements established before virtualization has become a mainstay of the IT. A backup policy is typically described as a principle or rule to guide decisions and achieve a rational outcome from some sort of interruption in the availability and integrity of protected data and systems.

A backup policy is defined as a set of rules to archive the required backup goals with the commonly associated terms such as: what, where, how, how much, how often and others. As this paper goes on, we'll help develop these terms as they apply to virtualization so you can craft your own policy.

At the highest level, the backup policy matches the requirements of common terms such as recovery point objective (RPO) and recovery time objective (RTO) with constraints like the backup window and budgetary parameters. A backup solution requires a design just as any other IT project process, which could be something similar to the figure below:

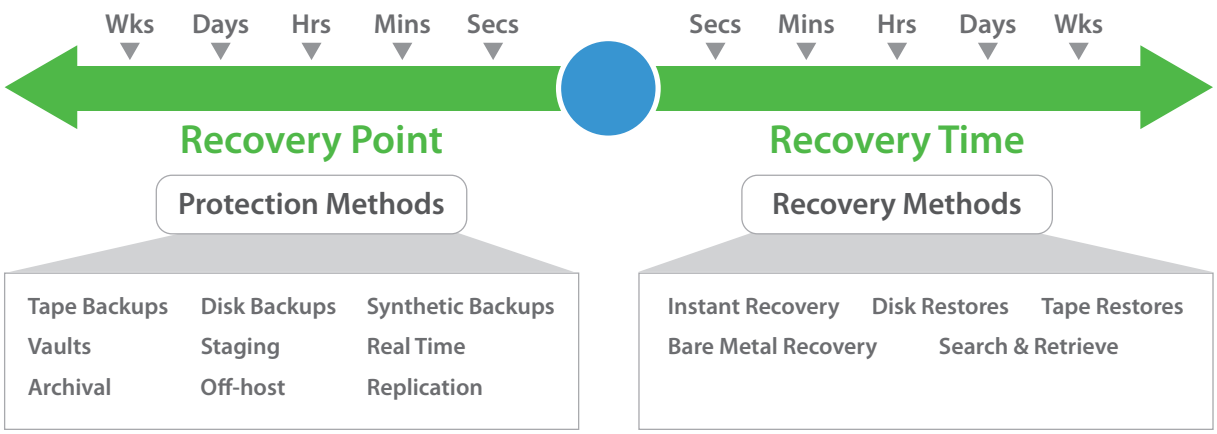


A backup policy usually does not specify the technical aspects that are related to specific backup solutions and products, much less address the influx of virtualization. If a backup solution is selected first, then a backup policy could become more specific and represent the capabilities of the product. Another approach would be to identify the specific requirements, then select a product which can deliver them.

In order to achieve the technical requirements and constraints of a backup policy, companies usually need a deep analysis to define the choices and identify any possible dependencies. Factors such as identifying a backup window can be a critical constraint and could conflict with some availability requirements.

Backup policies are part of larger plans related to business continuity initiatives, which are a business discussion that matches technical capabilities to the priorities of the organization. Various forms of data loss and important systems should be fully addressed with the RPO and RTO as defined in the organization's requirements. Be sure to check out video tracks on Backup Academy as they discuss RPO and RTO in many areas related to virtualization backups.

The figure below shows how some of the key aspects of a backup policy are aligned to technical aspects of a VMware infrastructure:



Backup policy: What will be backed up?

When any data protection strategy is discussed, the source object is critical. In the case of virtualized environments, the source is the VM that contains the data or applications that need to be protected. This source is not necessarily a property of a backup policy, but instead is just an input to a set of policies to define the backup job.

When we define the source of a backup, we must consider different aspects that may limit our choice or may create dependences within the backup strategy:

- **Required protection level:** A VM backup can be done in a number of different ways to achieve the desired RTO and RPO. A different level of protection may be needed for each VM.
- **How backups will be handled:** while usually a more technical aspect, this is a critical step for the overall backup architecture.
- **Type of sources:** Beyond VMs, we also need to be mindful of files, applications data, entire systems and other contents of the VMs that need to be protected.
- **Source size:** How big are the source VMs?

How to manage backup jobs and policies

The technical aspect, of course, depends on the backup products, but we can take a generic approach:

- **Agent or agent-less:** The backup program needs (or does not need) a specific agent on the source side (usually within the guest operating system) to handle the backup in the best possible way. This is common with application-level backups where the agent is used to manage the backup procedure inside the application.
- **Push or pull of the data:** How is the data flow managed, from the source or from the backup machine? Usually this aspect is not relevant, except in some special cases (geographic backup or backup across firewalls and NAT).

Source data types

This is an important aspect, because the backup set and the type of data on it can influence how backup (and also restore) procedures can be defined. The consistency of the backup data depends on the types of sources. A specific backup solution may not support all the types of sources, may require specific agents, options or components or may have different features available for virtual machines compared to physical systems, and vice versa.

The source object may include:

- **Files on the OS file system:** This is the simplest case, and usually at least a "crash consistent" backup can be guaranteed by most of the backup solutions (VSS1 or other solutions could be used to achieve better consistency). Although most backup solutions support files on Windows OS, not all can handle also files on Linux or other Unix OS.
- **Application data:** In this case, we need a deep knowledge of the type of application, should know how data is stored and how to handle the right consistency[1] of it, both for the backup and restore procedure. This usually requires backup agents for the specific applications and not all the backup solutions can handle the same set of applications.

[1] For more details, see the "Backing up enterprise applications: Transaction consistency is key" lesson on the Backup Academy.

- **Systems:** In this case, we want to protect the entire system and the related data instead of only a subset of its data. In order to have a consistent backup, we need not only files (and applications) consistency, but also system consistency (like the System State of a Windows Machine). By using specific features available in most of the backup products, we can perform, starting from a backup at the system level, a restore of the entire system or a restore at the file level (and in some cases, also a restore at the application level)[2]. On some backup products, the restore could be done also on a different system (that means at least reconfiguration of device drivers and boot loader) – also referred to as a bare metal recovery.
- **Virtual Machines:** This case is really similar to the previous one, except that now the systems are just virtual machines[3]. Depending on the type of hypervisor, we can handle backup more efficiently (compared to physical systems) and for some hypervisors[4] also in an agent-less way (for example, in VMware vSphere using VADP[5]). The backup is usually handled at the VM image level (the files that represent a VM) that is similar to a system image of a physical system; but, depending on the backup solutions, we can have a different level of restores (as describe in the previous point).

Usually file- and application-level backup can cover the data protection and recovery aspects; system and virtual machine backups (backup at the image level) are used to provide system protection and recovery. If the backup solution can handle file- and application-level restore from the system or VM image-level backup, it can ensure more flexibility and of course smaller backup windows.

Type of "transport" at source side

When we have chosen the data source, there could be different ways to handle the data flow from the source to the destination during a backup job:

- **Full copy:** on each job, the entire backup set is copied.
- **Incremental:** usually it is possible to track (on the source side) which objects are changed from the previous backup; in this way, we can only send this difference. Note that this could be simple for files (because the archive bit or the modified data can be used to find changed files), but more complicated for applications- and system-level backup (where usually an agent is needed to understand and handle the changes). For virtual environment, this could require specific functions (like the CBT[6] in VMware).

The right "transport" can be used to make backup more efficient and shrink the backup window (especially when the backup size is huge) by reducing the amount of data that must be transferred during a backup job.

Of course, in all cases a compression and/or deduplication (at the file or block level) at the source side could be used to reduce the amount of data, but this implies the need of an agent on the source and several resources to handle these activities.

On a shared storage environment and/or a virtual environment, the transport type may also define how data profiles are transferred during a backup operation: for example, using the common or a backup network (typical of most of the agent-based backup solutions), using the SAN (for example, by reading the data directly from the storage itself, instead of the sources)

---

[2] Some vendors call this function a "granular restore"

[3] For more details, there are several lessons on Backup Academy about the backup in virtual environment, for example: "Why virtual machine backups are different", "Core technologies used for virtual machine backup" and "Physical vs. Virtual Backups"

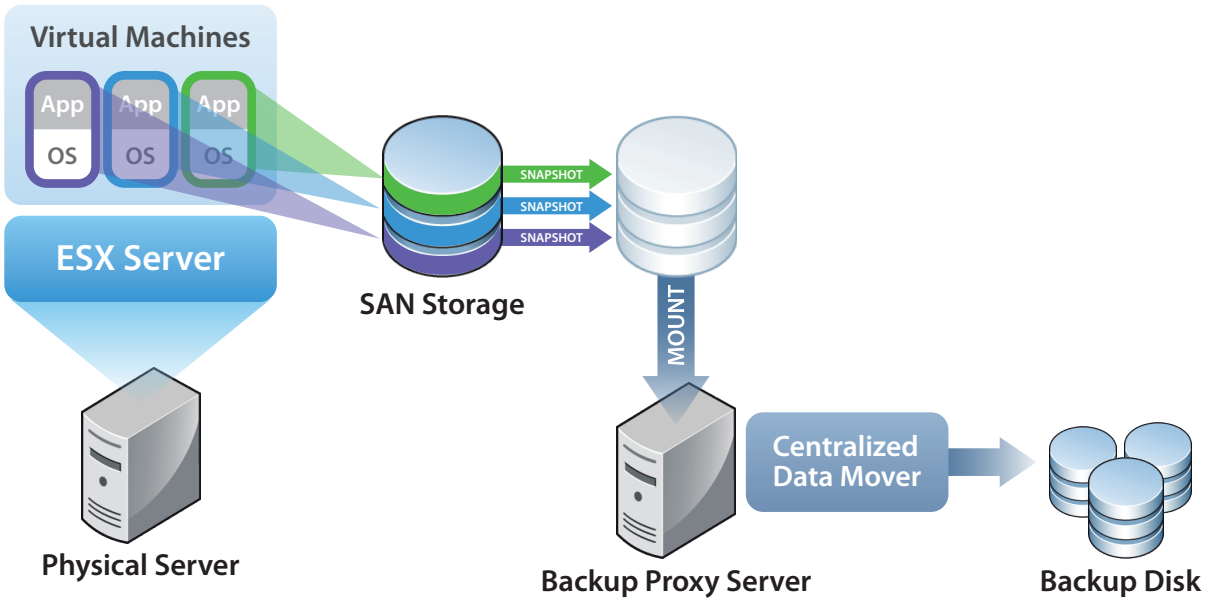
[4] For more details, see those lessons on Backup Academy: "Best practices for VMware backups" and "Best practices for Hyper-V backups"

[5] VMware KB: vStorage APIs for Data Protection (VADP) FAQ — http://kb.vmware.com/kb/1021175

[6] VMware KB: Changed Block Tracking (CBT) on virtual machines — http://kb.vmware.com/kb/1020128

or using other methods (like the hot-add transport mode on a VMware virtual environment). With a LAN-free solution, backup can be usually performed faster and/or with a less impact on the production environment.

In this picture, you can see an example of LAN-free backup of virtual machines, where VM snapshots are used and then those snapshots are mounted from the backup server directly from the SAN (on VMware vSphere, this special transport mode is referred to as SAN mode).



Size of the sources

Of course, this depends not only on the amount of data on it, but also on the type of the source; and an appropriate type of "transport" could be used to reduce this size (for example, by using incremental transfer). The type of source could increase the amount of data: for example, an image-level backup may include also unused space on the file system or deleted space that has not yet been reclaimed. How data change in the time may limit incremental transfer. And also keeping the right integrity and consistency may require more space.

The amount of data and how it is stored on the destination will also define some constrains on the type of destination and its size.

Size could also create some dependency on how to handle the consistency of the data: in most cases, some kinds of snapshot technologies are used and this means that additional space could be needed to handle the snapshot of large data profiles.

Backup policy: Where

Basically, we could have two main types of destinations:

- **Disks (D2D):** where backup data are stored on some kind of hard disk.
- **Tapes (D2T):** where backup data are stored on tapes.

Of course, there could be also other classifications of destinations: for example, using "cloud" approaches, we can have on-premise or off-premise. In a cloud approach, a destination could be a virtual datacenter or a part of it. Backup could become a service in the Backup as a Service (BaaS) model.

Also we can have some "hybrid" approach where both disks and tapes are used in a D2D2T way, where disks are used for "staging" (and then backups are moved to tape) or as the first level of backup. This kind of approach can gain the best of both worlds.

Disks

A backup to disk usually consists in some kind of structured files stored on:

- **Disks:** a set of disks, connected using DAS or SAN, usually organized in one or more logical volumes using some kind of RAID configuration (in most cases, RAID5 or RAID6 to maximize the available space but guarantee some kind of redundancy).
- **Network shares:** the destination is just a network share using standard NAS protocols (like NFS or CIFS/SMB). On the destination side, of course, there are some disks organized as in the previous point.
- **Some kind of hardware or virtual appliances:** usually they work as a NAS using network shares. But some may also work with iSCSI (by exporting a logical disk) or in some other way (for example, VTL, as explained in the next paragraph).

A disk-based solution has several pros:

- **Great capacity:** Tape capacity is limited (1.5 TB on LTO5) compared to disk capacity (2-3 TB on a single SATA disk, with many disk systems having multiple drives connected).
- **Speed:** Tape drive could be fast (140MBps uncompressed in a LTO5) compared to the average high-capacity SATA disk (around 35MBps), but disks can be "striped" together and also tape look-up and allocation requires a lot of time overhead. The restore time must also be considered and the lookup time is usually significant.
- **Scalability:** This is especially relevant in the appliance solutions where usually more appliances can be used together.
- **Replication and disaster recovery:** We can use storage-based replication or file-level replication to make remote copies of the backup data.
- **More flexible:** For example, deduplication technologies[7] could be possible, rather than a "simple" compression technology. Backup programs usually support at least backup to disk, yet not all also support the backup-to-tape approach.

Tapes

In the backup-to-tape solution, the destination usually consists in one kind of tape:

- **A tape unit:** That usually can handle only one tape and could be connected in a DAS way (internal or external using in most cases SAS or SCSI cables).
- **An autoloader:** That usually consists in an appliance with one or more tape units and several slots for tapes that could be connected using DAS or SAN.

[7] For more information about deduplication at destination level see the "Data deduplication in virtualized environments" lesson on Backup Academy

- **A tape library[8]:** Similar to an autoloader but bigger and usually with more size and features and using SAN connections.
- **Virtual Tape Library (VTL)[9]:** An appliance (hardware or software) that can be used as an autoloader or tape library but it works usually with disk instead of tapes.

Actually, tape standards are defined by the organization like the Linear Tape Open (LTO) consortium (whose member include Hewlett-Packard, IBM, and Quantum) that has officially released specifications for the LTO Generation 5 with around 1,5 TB of native capacity and athroughput of 140MBps (uncompressed)… and is working for the future standards.

Autoloaders and tape libraries are usually solutions which may be more expensive compared to similar disk-based solutions. One factor to consider is the cost per TB for tape compared to a SATA disk is becoming not so different. There is still a big advantage of a tape-based solution in that archiving a tape is not a so simple as it is with a disk. Though there are solutions using a removable cartridge that contains disks to simplify this process.

To gain the best of world, a D2D2T approach can be used, as described previously. It requires backup programs to support it, or more effort to the backup administrators to integrate and implement it.

There are also some cases there: the existing solution which is already based on backup to tape and a simple switch to a backup to disk approach is needed without changes on the backup side. In those cases, the VTL solutions with a disk-based appliance could be the right solution and could permit to use a native backup-to-disk solution in the future.

A VTL could also be an interesting approach to use SAN-based connections (like FC or iSCSI) to save the backup data, instead of NAS-based connections typical of most disk-based backup appliances.

[8] http://en.wikipedia.org/wiki/Tape_library
[9] http://en.wikipedia.org/wiki/Virtual_tape_library

Backup policy: How

This part of the policy will define several aspects:

- **Multi-tier:** How data is distributed using different types of destinations (as described in the previous chapter).
- **Destination "format":** How data is saved, in which format, with which kind of relations with previous data (full, incremental, differential, deduplicated, …).
- **Backup frequency:** How often are the backup jobs performed.
- **Backup retention:** How is old data removed from the destination and how much data must remain.

Multi-tier

We can use D2D and D2T solutions and combine them together to have two levels of backup in a D2D2T model, as described in the previous chapter.

But we can also have another level of backup: for example, the very first could be the storage itself, using (for example) storage snapshots. Formally, they are not a real backup but can be used to provide a really fast recovery procedure.

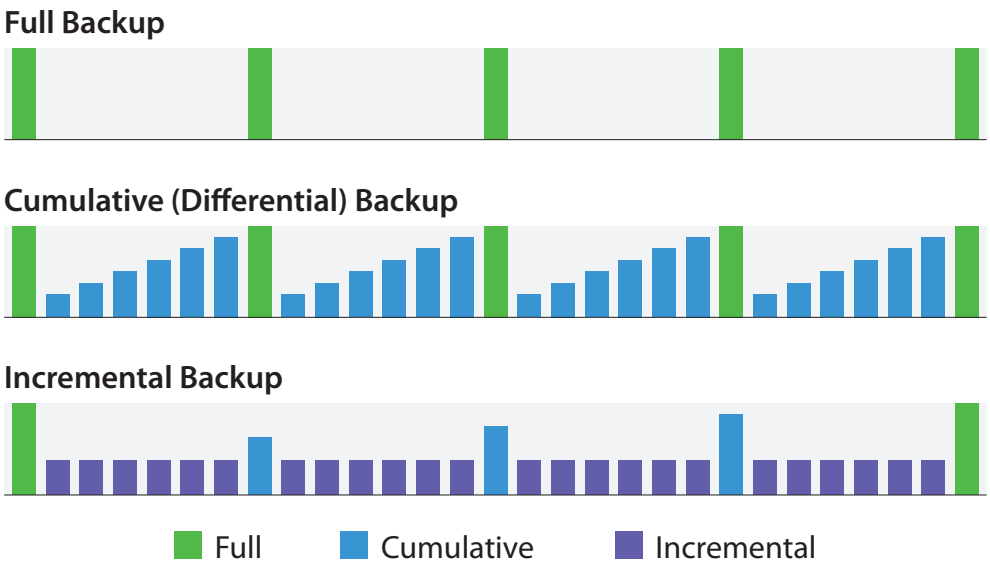
In an environment with replication, this could also be used as an additional level of backup, although there could be some limits (depending on the backup program) on replicas compared to real backups.

Destination "format"

The "format" of the backup data depends on the backup data, but it can also depend on the type of storage we are using: disk or tape. As described in the previous chapter, by using a D2D approach we can have different options to store the backup data. Usually, a set of files could be used, and a complete deduplicated archive could be implemented (this could be common on most of the D2D appliances).

On the tape destination, the type of format is quite limited: it usually consists of a sequence of compressed archives. Note that most tape units can handle the compression in hardware, so you can offload this process to the unit itself.

Depending on how the backup data is related with previous data, we can have at least these different formats:

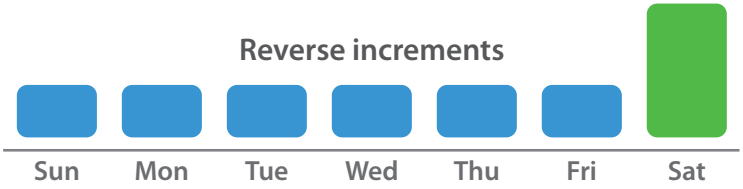


- **Full:** Each backup is an entire full backup of the source data. This does not imply that data must be transferred with a full backup from the source (except the first time), because the full set could also be rebuilt from the previous data and an incremental transfer.
- **Differential:** Each differential backup includes all data changes from the last full backup. Backup jobs are faster compared to the full backup and to make a restore, you just need a full data set and the last differential set.
- **Incremental:** Each incremental backup includes all data changes from the last backup. For this reason, backup jobs are usually faster, but to perform a restore, you may need (in the worst case) the last full backup and all incremental backups.
- **Additional technologies** such as synthetic full backups and changed block tracking can complement these core backup methods.

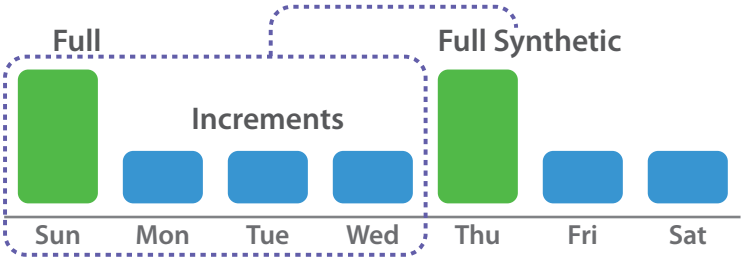
Those types of formats are normally used on D2T backup depending also on the retention of the data. Of course, they have some impact on the duration of backup and restore jobs.

In a D2D solution, another format could be implemented, like deduplication (as written in the previous chapter) as well as other variations of previous formats, like:

- **Reversed-increment:** the latest backup is always rebuilt as a full backup and the previous backups are incremental.



- **Synthetic backup:** that is an incremental model starting from a specific date and a reverse incremental before this date.



Note that security of backup data could be handled by some backup programs that use encryption. In case of D2D appliances, this may be performed by the appliance itself as some tape unit are able to offload this process.

Backup retention

Backup retention defines the minimum time for which data must be maintained and when old data can be deleted and/or purged. Of course, it depends on the previous properties, especially:

- The source type and its size (this is usually a requirement)
- The frequency (more backups could mean more data that could limit the retention)
- The "format" type (by using deduplication or incremental backup and/or compression, we can reduce the amount of data and increase the retention)
- The destination type (that could define the amount of data that could be stored)
- The required Recovery Point Objective (RPO)

Usually one could decide from the available destination space (a simple estimation if we are using compression and/or deduplication) and then define what could be a reasonable retention. But if the choice must be driven by the RPO, then the destination space must be designed to achieve the requirements.

In regards to retention, this topic could be handled in many different ways if we are performing backup to tape or backup to disk:

- D2D: This solution can usually scale with more disks to reach the required space, and incremental backup could be effective without major considerations in regard to the restore procedure. Also compression and/or deduplication can improve the available space. The reclaim of the space of expired data could be simple and fast (if backup data is stored on a file, this implies a simple file deletion).
- D2T: This solution uses a combination of full/differential/incremental models to improve the speed of the restore and try to store more information. But to scale with more space other policies are needed to change the tape media and put them off-line. Also the reclaim of the space could be more complicated and there are specific policies to handle the tape media rotation[10] (in order to recycle the old expired tape and also try to maximize the tape lifetime).

[10] http://en.wikipedia.org/wiki/Backup_rotation_scheme

Conclusion

A backup policy could be used to design and to document your backup technologies; but it can also be a springboard to more over-arching directives such as a BCP[11]. For this reason, my recommendation is to keep the set of backup policies at the high level and most vendor/solution independent as possible, in order to improve and validate the requirements as they apply to a specific product. Only in the final phase you can adapt and implement them on your specific solution, after the requirements are validated with a product.

Finally, although backup policies are usually applied to backup tasks, similar concepts could also be ported to other kind of solutions for data protection, such as replication and recovery validation. The process of managing a larger data protection strategy isn't just backups, it must include a number of options to address the larger requirements set forth by the organization.

For more information on this topic you can watch a Backup Academy session at http://www.backupacademy.com/basic-principles-backup-policies.html

[11] Business Continuity Plan